REMARKS

Claims 1-18 were pending in the application. Claims 1-18 stand rejected. Claims 1, 4-5, 10-11, and 14-15 were amended. Claims 19-26 were added. Claims 1-26 remain in the application.

Claim 10 stands rejected under 35 U.S.C. 101 as incorrectly reciting a computer program product. Claim 10 has been corrected.

Claims 5 and 15 stand rejected under 35 U.S.C. 112, first and second paragraphs. The rejection indicated that the term "color moment" did not appear in anywhere in the specification. Claims 5 and 15 have been amended to recite:

> "wherein the statistic is related to one or more of: color, texture, and straight lines in the digital image."

Claims 1-18 stand rejected under 35 U.S.C. 112, second paragraph. The rejection stated:

> 'In particular independent claims 1 and 11 use the word "semantic" to describe an object, which can be detected and from which, some kind of orientation can be gathered. The only definitions of semantic seemed to be directed to language or words.
>
> 'From dictionary.com - definition of Semantic:
>
> 1. Of or relating to meaning, especially meaning in language.
> 2. Of, relating to, or according to the science of semantics.
>
> 'In the environment of the present invention, semantic is being used to describe an object as a distinctive object from which orientation can be determined. So when applicant uses text as a semantic object the definition makes some sense, but when applicant uses for example in claim 3, human face, human figure, clear blue sky, lawn grass, a snow field, body of water, tree, a sign, and written text it is unclear how these objects are "semantic." Appropriate correction is required.'

The term "semantic", as used in the specification, does not disagree with a dictionary definition, particularly the emphasized language in the following:

> "semantic {*pronunciation, part of speech label, and derivation omitted*} 1 : of or relating to meaning in language <u>2 : of or relating to semantics</u>" *Webster's Ninth New Collegiate Dictionary*,

Merriam-Webster, Springfield, Massachusetts, (1990), page 1068. (emphasis added)

"semantics {*pronunciation, part of speech label, and derivation omitted*} 1 : the study of meanings: a : the historical and psychological study and the classification of changes in the signification of words or forms viewed as factors in linguistic development b (1) : SEMIOTIC (2) : a branch of semiotic dealing with the relations between signs and what they refer to and including theories of denotation, extension, naming, and truth 2 : GENERAL SEMANTICS 3 a : the meaning or relationship of meanings of a sign or set of signs; *esp* : connotative meaning b : the language used (as in advertising or political propaganda) to achieve a desired effect on an audience' (emphasis added)

The rejection makes the statement:

'In the environment of the present invention, semantic is being used to describe an object as a distinctive object from which orientation can be determined. So when applicant uses text as a semantic object the definition makes some sense, but when applicant uses for example in claim 3, human face, human figure, clear blue sky, lawn grass, a snow field, body of water, tree, a sign, and written text it is unclear how these objects are "semantic."

Applicants respectfully traverse. The rejection is arguing, quite literally, that the sky is not up. Each of the features mentioned tend to have inherent orientations, as is discussed in the application:

"Referring to Fig. 2, there is shown a typical consumer snapshot photograph. This photo contains a plurality of notable semantic objects, including a person with a human face region **100**, a tree with a tree crown (foliage) region **101** and a tree trunk region **110**, a white cloud region **102**, a clear blue sky region **103**, a grass region **104**, a park sign **107**, and other background regions. Many of these semantic objects have unique upright orientation by themselves and their orientations are often correlated with the correct orientation of the entire image (scene). For example, people, trees, text, signs are often in upright positions in an image, sky and cloud are at the top of the image, while grass regions **104**, snow fields (not shown), and open water bodies such as river, lake, or

ocean (not shown) tend to be at the bottom of an image." (application, page 8, lines 4-14; emphasis added)

Each of human face, human figure, clear blue sky, lawn grass, a snow field, body of water, tree, a sign, and written text meets the language: "semantic is being used to describe an object as a distinctive object from which orientation can be determined."

Claims 1-5, 8-15, and 18 stand rejected under 35 U.S.C. 102(e) as being anticipated by U.S. Patent 6,798,905 to Sugiura et al. The rejection stated:

> 'With regard to claim 1, Sugiura discloses a method for determining the orientation of a digital image, comprising the steps of:
>
> 'a) employing a semantic object detection method to detect the presence and orientation of a semantic object (column 9, lines 20-25). Sugiura discloses using text as a semantic object. The orientation of the text is calculated and then a "line direction" or orientation is determined.
>
> 'b) employing a scene layout detection method to detect the orientation of a scene layout (column 9, lines 25-30). The scene layout detection is interpreted as the use of line direction to determine the orientation for using each sub-image.
>
> 'c) employing an arbitration method to produce an estimate of the image orientation from the orientation of the detected semantic object and the detected orientation of the scene layout (column 9, lines 30-35). Sugiura discloses that each sub-image has an orientation (scene layout) calculated by using text and line direction information (semantic object and orientation). Then each calculated orientation has a reliability measure calculated and the one with the greatest reliability measure is the calculated orientation to be used. So the arbitration uses information from both the semantic object orientation and the scene layout orientation (sub-image orientation).'

Claim 1 has been amended to state:

> 1. A method for determining the orientation of a captured digital image, comprising the steps of:
>
> a) employing a semantic object detection method to detect the presence and orientation of a semantic object in the digital image;

b) employing a scene layout detection method to detect the orientation of a scene layout of the digital image; and

c) employing an arbitration method to produce an estimate of the image orientation by arbitrating between the orientation of the detected semantic object and the detected orientation of the scene layout.

Claim 1 is supported by the application as filed, notably the original claims and at page 8, lines 4-5; page 12, lines 2-6; page 13, line 27 to page 14, line 2.

Claim 1 requires that the digital image is captured and that the respective detection methods detect a semantic object orientation in the digital image and orientation of <u>a scene layout of the digital image</u>. As the rejection indicates, "scene layout" in Sugiura is not of a digital image, but rather of sub-images:

> "Sugiura discloses that each sub-image has an orientation (scene layout) ... So the arbitration uses information from both the semantic object orientation and the scene layout orientation (sub-image orientation)." (page 5; see also Sugiura, col. 8, line 66 to col. 9, line 9; col. 9, lines 17-26)

The scene layout orientation of Sugiura referred to in the rejection is, thus, unlike that of Claim 1.

Claim 1 requires employing an arbitration method to produce an estimate of the image orientation by arbitrating between the orientation of the detected semantic object and the detected orientation of the scene layout of the digital image, not an area of digital image. In Sugiura, there is a detection of the orientation of the digital image, but that detection is performed by an orientation detection process, <u>only after a particular area is selected using a reliability measure</u>. (Sugiura, col. 10, lines 28-40)

There is also no arbitrating in Sugiura between the orientation of a semantic object, which is identified in the rejection as:

> 'Sugiura discloses using text as a semantic object. The orientation of the text is calculated and then a "line direction" or orientation is determined.' (office action, page 4)

and the orientation of a scene layout of the digital image, which is identified in the rejection as:

"The scene layout detection is interpreted as the use of line direction to determine the orientation for using each sub-image." (office action, page 4))

These are both the same, at least for the purposes of arbitrating between them. The line direction of an area is the orientation of the text, which is the orientation of the area. The rejection identified Sugiura, column 9, lines 30-35 with employing an arbitration method, but that portion of Sugiura deals only with determination of a reliability measure of each the areas having a line direction. Even if that portion were, for the sake of argument, to be considered arbitrating between the orientations of the different areas, such arbitrating would not meet the language of Claim 1, which requires arbitrating between the orientation of the detected semantic object and the detected orientation of the scene layout.

Claims 2-5 and 8-10 are allowable as depending from Claim 1 and as follows.

The rejection stated in relation to Claim 4:

'With regard to claim 4, Sugiura discloses therein the scene layout detection method comprises the steps of:

'a) dividing the digital image into non-overlapping blocks (Fig. 6) ;

'b) computing at least one statistic for each image block (column 9, lines 26-30);

'c) forming a feature vector by concatenating the statistics computed from the image blocks (column 9, lines 26-30). Sugiura discloses obtaining a reliability measure for each sub-image block. The reliability measure is calculated from the corresponding histograms. The histograms are interpreted as a feature vector.

'd) using a trained classifier to produce an estimate of the image orientation (column 9, lines 25-35). The trained classifier producing an estimate are the "estimates" or orientations calculated for each sub-image, based on these "estimates" reliability measures are calculated and the final orientation is decided.'

Claim 4 was amended to state:

4. The method claimed in claim 1, wherein the scene layout detection method comprises the steps of:

a) dividing the digital image into non-overlapping image blocks;

b) computing at least one statistic for each image block;

c) forming a feature vector by concatenating the statistics computed from the image blocks; and

d) using a classifier trained with a plurality of scene prototype images to produce an estimate of the image orientation.

Amended Claim 4 is supported by the application as filed, notably the original claims and at page 8, line 23 to page 9, line 8.

Claim 4 requires using a classifier trained with a plurality of scene prototypes to produce an estimate of the image orientation. This feature is discussed at length in the application:

> 'Referring to Fig. 3, one possible embodiment of the spatial layout detector 210 will be described. A collection of training images 300, preferably those that fall into scene prototypes, such as "sunset", "beach", "fields", "cityscape", and "desert", are provided to train a classifier 340 through learning by example. Typically, a given image is partitioned into small sections. A set of characteristics, which may include color, texture, curves, lines, or any combination of these characteristics are computed for each of the sections. <u>These characteristics, along with their corresponding positions, are used as features that feed the classifier.</u> This process is referred to as feature extraction 310. Using a statistical learning procedure 320 (such as described in the textbook: Duda, et al., "Pattern Classification", John Wiley & Sons, 2001), parameters 330 of a suitable classifier 340, such as a support vector machine or a neural network, are obtained. In the case of a neural network, the parameters are weights linking the nodes in the network. In the case of a support vector machine, the parameters are the support vectors that define the decision boundaries between different classes (in this case, the four possible orientations of a rectangular image) in the feature space. <u>This process is referred to as "training". The result of the training is that the classifier 340 learns to recognize scene prototypes that have been presented to it during training.</u> One such prototype is shown in Fig. 4, which can be categorized as "blue color and no texture at the top 500, green color and light texture at the

bottom 510". For a test image 301, usually not part of the training images, the same feature extraction procedure 310 described above is applied to the test image to obtain a set of features. Based on values of these features, the trained classifier 340 would find the closest prototype and produce an estimate of the image orientation 350 based on the orientation of the closest matched prototype. For example, the prototype shown in Fig. 4 would be found to best match the image shown in Fig. 2. Therefore, it can be inferred that the image is already in the upright orientation.' (application, page 8, line 15 to page 9, line 8; emphasis added)

The rejection proposes that Sugiura has 'The trained classifier producing an estimate are the "estimates" or orientations calculated for each sub-image', but this proposed feature lacks a "plurality of scene prototype images". It will be noted, that the application in the above quote, defines "training". The Sugiura "classifier" is not trained by this definition. For example, lacking a plurality of scene prototype images, the Sugiura "classifier" cannot have leaned "to recognize scene prototypes that have been presented to it during training" (see above quote).

The rejection stated in relation to Claim 5:

"With regard to claim 5, Sugiura discloses that the invention is applied to full-color copiers. It is unclear what a color moment is because there is no support or definition in the present specification, but clearly the invention of Sugiura is in the environment of full-color imaging."

Claim 5 was amended to state:

5. The method claimed in claim 4, wherein the statistic is related to one or more of: color, texture, and straight lines in the digital image.

Amended Claim 5 is supported by the application as filed, notably the original claims and at page 7, lines 15-18.

The rejection stated in relation to Claim 11:

"With regard to claim 11, Sugiura discloses a system for performing the method steps as discussed in claim 1 (Figs. 2, 4 and 5).

Claim 11 was amended to state:

11. A system for processing a digital color image, comprising:

a semantic object detector to determine the presence and orientation of a semantic object in the digital color image;

a scene layout detector to determine the orientation of a scene layout of the digital color image, said scene layout detector having a classifier trained with a plurality of scene prototype images;

an arbitrator responsive to the orientation of the semantic object and the orientation of the scene layout to produce an estimate of the image orientation; and

an image rotator to re-orient the digital image in the upright direction.

Claim 11 is supported and allowable on the grounds discussed above in relation to Claim 4.

Claims 12-15 and 18 are allowable as depending from Claim 11 and as follows.

Claim 15 is supported and allowable on the same grounds as Claim 5.

Claims 6 and 16 stand rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent 6,798,905 to Sugiura et al. in view of IEEE document titled "Vanishing Point Detection by Line Clustering" by G.F. McLean and D. Kotturi hereinafter referred to as McLean. Claims 6 and 16 are allowable as depending from Claims 1 and 11, respectively.

Claims 7 and 17 stand rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent 6,798,905 to Sugiura et al. in view of U.S. Patent 6,996,549 to Zhang et al. Claims 7 and 17 are allowable as depending from Claims 1 and 11, respectively.

Added Claim 19 states:

"19. The method of claim 1 wherein said captured digital image is of a natural scene."

Claim 19 is supported by the application as filed, notably page 6, line 19 and page 8, lines 4-14. Claim 19 is allowable as depending from Claim 1 and as follows. Sugiura requires an image of a textual document, since it orients using line direction determined from text and assumes that a most reliable text line

direction is the orientation of the image. (See citations to Sugiura in the above discussion of Claim 1.) As the rejection notes and as discussed above:

> 'Sugiura discloses using text as a semantic object. The orientation of the text is calculated and then a "line direction" or orientation is determined.' (office action, page 4)
>
> "The scene layout detection is interpreted as the use of line direction to determine the orientation for using each sub-image." (office action, page 4)

This is incompatible with Claim 19, which requires:

> a) employing a semantic object detection method to detect the presence and <u>orientation of a semantic object in the digital image of a natural scene</u>;
>
> b) employing a scene layout detection method to detect the <u>orientation of a scene layout of the digital image of a natural scene</u>; and
>
> c) employing an arbitration method to produce an estimate of the image orientation by arbitrating between the orientation of the detected semantic object and the detected orientation of the scene layout.

Added Claim 20 states:

> 20. The method of claim 19 wherein the orientation of the detected semantic object and the detected orientation of the scene layout are contradictory.

Claim 20 is supported by the application as filed, notably the original claims and at page 12. lines 5-6. This is unlike Sugiura, since it is not apparent how a line direction orientation and an area orientation (the same line direction) can be contradictory. (See the above discussion of Claim 1.)

Added Claim 21 states:

> 21. The system of claim 2, wherein:
>
> each of said semantic object detectors detects the respective said semantic object orientation to be any one of: upright orientation, upside-down orientation, left-to-right orientation, right-to-left orientation, and undecided orientation and
>
> said scene layout detection method determines the orientation of the scene layout to be any one of: upright orientation,

upside-down orientation, left-to-right orientation, and right-to-left orientation.

Claim 21 is supported by the application as filed, notably the original claims and at page 11, lines 7-27 and page 1, lines 18-28; page 8, lines 28-30.

Claim 21 is allowable as depending from Claim 2 and as follows. Claim 21 requires that each of the semantic object orientations detected is any one of upright, upside-down, left-to-right, right-to-left, and undecided and the scene layout orientation is any one of upright orientation, upside-down orientation, left-to-right orientation, and right-to-left orientation. This is not possible in Sugiura. The line direction orientation (interpreted as the semantic object orientation by the office action) is used for the area orientation (interpreted as the scene layout orientation by the office action). (Sugiura, col. 9, lines 26-30) Sugiura does disclose determining the degree of reliability of an area orientation, but that determination is based on the area orientations:

> "the reliability judging unit 240 next judges reliability for each area in accordance with the histogram of the line direction of that area." (Sugiura, col. 9, lines 26-30)

If the reliability is low, Sugiura teaches not using that area's orientation. (Sugiura, col. 10, lines 4-7) This contrasts with Claim 21, which since it depends from Claims 1 and 2, requires that the orientation of the detected semantic object, any one of upright, upside-down, left-to-right, right-to-left, and undecided, is one of the orientations that is arbitrated between.

Added Claim 22 states:

> 22. A system for processing a digital image comprising:
>
> one or more semantic object detectors, each said semantic object detector being adapted to determine the presence and orientation in the digital image of a semantic object of a respective one of a plurality of different types;
>
> a scene layout detector adapted to determine the orientation of a scene layout of the digital image; and
>
> an arbitrator adapted to arbitrate between said determined semantic object and scene layout orientations and produce an estimate of an orientation of the digital image.

Claim 22 is supported and allowable on grounds discussed above in relation to Claim 1.

Added Claims 23-26 are allowable as depending from Claim 22 and as follows.

Added Claim 23 states:

> 23. The system of claim 22, wherein each of said semantic object detectors determines the respective said semantic object orientation to be any one of: upright orientation, upside-down orientation, left-to-right orientation, right-to-left orientation, and undecided orientation.

Claim 23 is supported by the application as filed, notably the original claims and at page 11, lines 7-27 and page 1, lines 18-28; page 8, lines 28-30.

Claim 23 is allowable as depending from Claim 22 and as follows. Claim 23 requires that the orientation of the detected semantic object, any one of upright, upside-down, left-to-right, right-to-left, and undecided, is one of the orientations that is arbitrated between. This is not taught or suggested by Sugiura, as discussed above in relation to Claim 21.

Claim 24 is allowable as depending from Claim 23 and is supported and allowable in the same manner as Claim 20.

Claim 25 states:

> 25. The method of claim 22 wherein said scene layout detector has a classifier trained with a plurality of scene prototype images.

Claim 25 is supported and allowable on grounds discussed above in relation to Claim 4.

Claim 26 states:

> 26. The method of claim 25, wherein said classifier is one of a support vector machine and a neural network.

Claim 26 is supported by the application as filed, notably at page 8, lines 23-26. Claim 26 is allowable as depending from Claim 25 and as requiring the additional feature of the classifier being of a specific type.

It is believed that these changes now make the claims clear and definite and, if there are any problems with these changes, Applicants' attorney would appreciate a telephone call.

In view of the foregoing, it is believed none of the references, taken singly or in combination, disclose the claimed invention. Accordingly, this application is believed to be in condition for allowance, the notice of which is respectfully requested.

Respectfully submitted,

Attorney for Applicant(s)
Registration No. 30,700

Robert Luke Walker/amb
Rochester, NY 14650
Telephone: (585) 588-2739
Facsimile: (585) 477-1148

Enclosures: Transmittal Fee Sheet (original and copy)
Request for Three Month Extension of Time (original and copy)
PTO-2038 (original & copy)
Postcard